LARGE DEVIATIONS APPLICATION TO BILLINGSLEY'S EXAMPLE

R. LIPTSER

Abstract. We consider a classical model related to an empirical distribution function $F_n(t)=\frac{1}{n}\sum_{k=1}^n I_{\{\xi_k\le t\}}$ of $(\xi_k)_{i\ge 1}$ – i.i.d. sequence of random variables, supported on the interval $[0,1]$, with continuous distribution function $F(t)=\mathsf{P}(\xi_1\le t)$. Applying "Stopping Time Techniques", we give a proof of Kolmogorov's exponential bound

$$\mathsf{P}\big(\sup_{t\in[0,1]}|F_n(t)-F(t)|\ge\varepsilon\big)\le \text{const}.e^{-n\delta_\varepsilon}$$

conjectured by Kolmogorov in 1943. Using this bound we establish a best possible logarithmic asymptotic of

$$\mathsf{P}\big(\sup_{t\in[0,1]}n^\alpha|F_n(t)-F(t)|\ge\varepsilon\big)$$

with rate $\frac{1}{n^{1-2\alpha}}$ slower than $\frac{1}{n}$ for any $\alpha\in\big(0,\frac{1}{2}\big)$.

1. Introduction

Let $(\xi_k)_{i\ge 1}$ be the i.i.d. sequence of random variables with values in the interval $[0,1]$ having a continuous distribution function $F(t)=\mathsf{P}(\xi_1\le t)$. Consider an empirical distribution $F_n(t)=\frac{1}{n}\sum\limits_{k=1}^n I_{\{\xi_k\le t\}}$. A strong law of large numbers for sums of i.i.d. random variables guaranties that for any $t\in[0,1]$, $F_n(t)\xrightarrow[n\to\infty]{\text{a.s.}}F(t)$ and the Glivenko-Cantelli theorem also guarantees a uniform convergence $\sup_{t\in[0,1]}|F_n(t)-F(t)|\xrightarrow[n\to\infty]{\text{a.s.}}0$.

For any fixed t, the rate of convergence, in $n\to\infty$, of $|F_n(t)-F(t)|$ is also well known from the Central Limit Theorem (CLT): $\{\frac{1}{\sqrt{n}}[F_n(t)-F(t)]\}_{n\to\infty}$ converges in law to a zero mean Gaussian random variable with the variance $F(t)[1-F(t)]$.

From Theorem 16.4 of Billingsley (1968), it is known that the family $\big\{\big(\frac{1}{\sqrt{n}}[F_n(t)-F(t)]\big)_{t\in[0,1]}\big\}_{n\to\infty}$ converges in law (in Skorokhod's and uniform metrics) to a zero mean Markov-Gaussian process $X=(X_t)_{t\in[0,1]}$ with a correlation function

$$K(t,s)=F(s\wedge t)[1-F(s\vee t)].\tag{1.1}$$

For $F(t)\equiv t$, the limit random process X is known as "Brownian Bridge" defined as the unique solution of Itô's equation $X_t=-\int_0^t\frac{X_s}{1-s}ds+B_t$ relative to Brownian motion B_t. In the general case, $F(t)\not\equiv t$, the random process

1991 *Mathematics Subject Classification.* 60F10, 60J27.
Key words and phrases. Empirical distribution; LDP; Stopping time.

X can be also defined as the unique solution of Itô's equation

$$X_t = -\int_0^t \frac{X_s}{1-F(s)} dF(s) + \mathbf{M}_t \tag{1.2}$$

with Brownian motion B_t replaced by a Gaussian martingale $\mathbf{M}_t$, $\mathsf{E}\mathbf{M}_t^2 \equiv F(t)$ (see Section 2.1).

Once, Prof. A.N. Shiryaev has mentioned to participants of the Probability Seminar at the Steklov Mathematical Institute that in 1943 Kolmogorov conjectured the following rate of convergence in the uniform metric,

$$\mathsf{P}\Big(\sup_{t\in[0,1]} |F_n(t) - F(t)| \ge \varepsilon\Big) \le \text{const.} e^{-n\delta_\varepsilon}, \tag{1.3}$$

a proof of which has never been published.

In this paper, we give a version of Kolmogorov's exponential bound with

$$\delta_\varepsilon = \frac{\varepsilon}{8}\Big\{\log\Big(1+\frac{\varepsilon^2}{32}\Big) - 1\Big\} + \frac{4}{\varepsilon}\log\Big(1+\frac{\varepsilon^2}{32}\Big).$$

It should be noted that neither Sanov's theorem (1961), [9] (see also Dembo Zeitouni, [2]) nor Wu's result (1994), [12], are not relevant tools for obtaining the Kolmogorov bound (1.3), since the Levy-Prohorov metric is involved in Sanov (1961) and Wu (1994). A crucial role in proving of (1.3) plays "Stopping Time Techniques".

Unfortunately, we could not claim that (1.3) is best possible bound even in a logarithmic scale. However, the Kolmogorov bound helps us to establish the following logarithmic asymptotics: for any $\alpha \in \big(0, \frac{1}{2}\big)$ and any T in a small vicinity of $\{1\}$,

$$\lim_{n\to\infty} \frac{1}{n^{1-2\alpha}} \log \mathsf{P}\Big(\sup_{t\in[0,T]} n^\alpha |F_n(t) - F(t)| \ge \varepsilon\Big) = -2\varepsilon^2. \tag{1.4}$$

We build the proof of (1.4) based on Kolmogorov's bound and on a non-standard Large Deviations technique. A key for (1.4) consists in choosing the rate $\frac{1}{n^{1-\alpha}}$ slower than $\frac{1}{n}$.

The paper is organized as follows. Section 2 contains auxiliary results from Stochastic Calculus useful for the asymptotic analysis of the random process $\big(F_n(t) - F(t)\big)_{t\in[0,T]}$ as $n \to \infty$. Proofs of (1.3) and (1.4) are given in Sections 3 and 4 respectively. The Large Deviations Principle result is formulated and proved in Section A (Appendix).

2. Stochastic calculus applicability to Billingsley's theorem

2.1. X_t as a Solution of (1.2). The limit random processes $X = (X_t)_{t\in[0,1]}$ is zero mean Gaussian with the correlation function defined in (1.1). By Theorem 8.1 of Doob (1953), the gaussianity of X jointly with an obvious property of the correlation function,

$$K(t,s) = \frac{K(t,u)K(u,s)}{K(u,u)},$$

enable us to claim that X is Markov process with respect to a minimal filtration $(\mathscr{F}_t^X)_{t\in[0,1]}$ generated by X. Then for $s<u<t$,

$$\mathsf{E}\Big(\frac{X_t}{1-F(t)}\Big|\mathscr{F}_u^X\Big)=\mathsf{E}\Big(\frac{X_t}{1-F(t)}\Big|X_u\Big)=\frac{1}{1-F(t)}\frac{K(t,u)}{K(u,u)}X_u=\frac{X_u}{1-F(u)}.$$

In other words, the Gaussian random process $N_t=\frac{X_t}{1-F(t)}$ is the square integrable martingale, i.e., a process with orthogonal increments (so, with independent increments too). Hence, its predictable variation process $\langle N\rangle_t$ coincides with $\mathsf{E}N_t^2=\frac{K(t,t)}{[1-F(t)]^2}=\frac{F(t)}{1-F(t)}$.

Therefore, the process $\mathbf{M}_t=\int_0^t[1-F(s)]dN_s$ is the Gaussian martingale with

$$\langle M\rangle_t=\int_0^t[1-F(s)]^2d\langle N\rangle_s=\int_0^t[1-F(s)]^2d\Big(\frac{F(s)}{1-F(s)}\Big)=F(t).$$

Finally, the Itô equation (1.2) is derived by applying the Itô formula to $X_t=[1-F(t)]N_t$.

2.2. **Counting Process $\sum\limits_{k=1}^{n} I_{\{\xi_k\le t\}}$.** Without loss of generality we shall assume that all ξ_k's are defined on a probability space $(\Omega,\mathcal{F},\mathsf{P})$. Denote

- $\mathscr{F}^k=(\mathscr{F}_t^k)_{0\le t\le 1}$ the filtration generated by $I_{\{\xi_k<t\}}$,
- $\mathscr{F}_t=\bigvee_{k\ge 1}\mathscr{F}_t^k$,
- $\mathcal{F}=\bigvee_{t\in[0,1]}\mathscr{F}_t$

and assume also that the general conditions for these filtrations are fulfilled.

The random process $I_{\{\xi_k\le t\}}$ has piece-wise constant and right continuous paths with only one jump of the unit size. Thus, $(I_{\{\xi_k\le t\}},\mathscr{F}_t^k)_{t\in[0,1]}$ is a counting process with continuous (!) compensator $(A_t^k)_{t\in[0,1]}$,

$$A_t^k=\int_0^{t\wedge\xi_k}\frac{dF(s)}{1-F(s)}=\int_0^t\frac{1-I_{\{\xi_k\le s\}}}{1-F(s)}dF(s)$$

(see, e.g., formula (18.23), Section 18.2 in [6]). Set $M_t^k=I_{\{\xi_k\le t\}}-A_t^k$. It is well known (see, e.g., Ch. 18 in [6]) that $(M_t^k,\mathscr{F}_t^k)_{t\in[0,1]}$ is a square integrable martingale with paths from the Skorokhod space $\mathbb{D}_{[0,1]}$ and its predictable quadratic variation process $\langle M^k\rangle_t\equiv A_t^k$. The joint independence of $(\xi_k)_{k\ge1}$ implies that $\{(I_{\{\xi_k\le t\}},\mathscr{F}_t)_{t\in[0,1]}\}_{k\ge1}$ are counting processes with disjoint jumps. Set $\mathbf{I}_t^n=\sum\limits_{k=1}^n I_{\{\xi_k\le t\}}$. Then, $(\mathbf{I}_t^n,\mathscr{F}_t)_{t\in[0,1]}$ is a counting process with the corresponding compensator,

$$\mathbf{A}_t^n=\sum_{k=1}^n A_t^k=n\int_0^t\frac{1-F_n(s)}{1-F(s)}dF(s)\tag{2.1}$$

or, equivalently, $\big(\mathbf{I}_t^n-\mathbf{A}_t^n,\mathscr{F}_t\big)_{t\in[0,1]}$ is the square integrable martingale with the predictable variation process $\mathbf{A}_t^n$. Two other martingales are related to $\big(\mathbf{I}_t^n-\mathbf{A}_t^n,\mathscr{F}_t\big)_{t\in[0,1]}$: $(\mathbf{M}_t^n,\mathscr{F}_t)_{t\in[0,1]}$ and $(\mathbf{M}_t^{n,\alpha},\mathscr{F}_t)_{t\in[0,1]}$, where

$$\mathbf{M}_t^n=\frac{1}{\sqrt{n}}\big(\mathbf{I}_t^n-\mathbf{A}_t^n\big)\quad\text{and}\quad\mathbf{M}_t^{n,\alpha}=\frac{1}{n^{\frac12-\alpha}}\mathbf{M}_t^n,\ \alpha\in\Big[0,\frac12\Big),$$

with predictable variation processes respectively:

$$\langle \mathbf{M}^n\rangle = \frac{1}{n}\mathbf{A}^n_t \quad \text{and} \quad \langle \mathbf{M}^{n,\alpha}\rangle_t = \frac{1}{n^{2(1-\alpha)}}\mathbf{A}^n_t.$$

2.3. **Functional Central Limit Theorem for $\mathbf{M}^n_t$.**

Theorem 2.1. *The family of martingales* $\big\{\big(M^n_t\big)_{t\in[0,1]}\big)\big\}_{n\to\infty}$ *converges in law* (*in Skorokhod's and uniform metrics*) *to a Gaussian martingale* $\mathbf{M}_t$ *with* $\langle \mathbf{M}\rangle_t = F(t)$.

Proof. In view of the function $F(t)$ is continuous, the Gaussian martingale $\mathbf{M}_t$ is continuous too. Then, by Theorem 2, Ch. 7, §1 of Liptser-Shiryaev (1989), [5], the desired statement holds true provided that $\big\langle \mathbf{M}^n\big\rangle_t \xrightarrow[n\to\infty]{\text{prob.}} F(t)$, $\forall\ t\in[0,1]$. The latter holds since

$$\big\langle \mathbf{M}^n\big\rangle_t = \frac{1}{n}\mathbf{A}^n_t = \frac{1}{n}\sum_{k=1}^n\int_0^t \frac{1-I_{\{\xi_k\le s\}}}{1-F(s)}dF(s)$$

and, in the case under consideration, the strong law of large numbers for sums of i.i.d. random variables implies

$$\lim_{n\to\infty}\Big\langle \mathbf{M}^n\Big\rangle_t = \mathsf{E}\int_0^t \frac{1-I_{\{\xi_1\le s\}}}{1-F(s)}dF(s) = F(t) \text{ a.s. } \forall\ t\in[0,1].$$

$\square$

2.4. **Semimartingale Decomposition of Centered Empirical Distribution.** Set

$$X^{n,\alpha}_t = n^\alpha\big[F_n(t)-F(t)\big],\ \alpha\in\Big[0,\frac{1}{2}\Big]. \tag{2.2}$$

Lemma 2.1. *For* $t\in[0,1)$,

- **(i)** $X^{n,\alpha}_t = -\int_0^t \frac{X^{n,\alpha}_s}{1-F(s)}dF(s) + \frac{1}{n^{\frac12-\alpha}}\mathbf{M}^n_t$;
- **(ii)** $X^{n,\alpha}_t = \frac{1}{n^{\frac12-\alpha}}[1-F(t)]\int_0^t \frac{d\mathbf{M}^n_s}{1-F(s)}$;
- **(iii)** $X^{n,\alpha}_t = \frac{1}{n^{\frac12-\alpha}}\Big\{\mathbf{M}^n_t - [1-F(t)]\int_0^t \frac{\mathbf{M}^n_s}{[1-F(s)]^2}dF(s)\Big\}$;
- **(iv)** $X^{n,\alpha}_t = \Psi\Big(\frac{1}{n^{\frac12-\alpha}}\mathbf{M}^n_{[0,t]}\Big)$, *where for any function* $(x_t)_{t\in[0,1]}$ *from the Skorokhod space* $\mathbb{D}_{[0,1]}$,

$$\Psi\big(x_{[0,t]}\big) = x_t - [1-F(t)]\int_0^t \frac{x_s}{[1-F(s)]^2}dF(s)$$

is continuous function in the uniform metric on $[0,1]$.

Proof. **(i)** From (2.2) and the definition of $\mathbf{A}_t^n$ and $\mathbf{M}_t^n$, it follows that $X_t^{n,\alpha} = \frac{1}{n^{1-\alpha}} \sum\limits_{k=1}^{n} [A_t^k - F(t)] + \frac{1}{n^{\frac{1}{2}-\alpha}} \mathbf{M}_t^n$. Consequently,

$$
\begin{aligned}
&\frac{1}{n^{1-\alpha}} \sum_{k=1}^{n} [A_t^k - F(t)] \\
&= \frac{1}{n^{1-\alpha}} \sum_{k=1}^{n} \left[\int_0^{t \wedge \xi_k} \frac{dF(s)}{1-F(s)} - F(t) \right] \\
&= \frac{1}{n^{1-\alpha}} \sum_{k=1}^{n} \left[\int_0^{t} \frac{1 - I_{\{\xi_k \le s\}}}{1-F(s)} dF(s) - F(t) \right] \\
&= -\int_0^t \frac{n^\alpha [F_n(s) - F(s)]}{1-F(s)} dF(s) = -\int_0^t \frac{X_s^{n,\alpha}}{1-F(s)} dF(s).
\end{aligned}
$$

(ii) This formula describes the unique solution of Itô's equation from **(ii)**

(iii) The Itô formula $\frac{\mathbf{M}_t^n}{1-F(t)} = \int_0^t \frac{d\mathbf{M}_s^n}{1-F(s)} + \int_0^t \frac{\mathbf{M}_s^n}{[1-F(s)]^2} dF(s)$ and **(ii)** provide

$$
\begin{aligned}
\frac{1}{n^{\frac{1}{2}-\alpha}} \mathbf{M}_t^n &= \frac{1}{n^{\frac{1}{2}-\alpha}} [1-F(t)] \left\{ \frac{\mathbf{M}_t^n}{1-F(t)} \right\} \\
&= \frac{1}{n^{\frac{1}{2}-\alpha}} [1-F(t)] \left\{ \int_0^t \frac{d\mathbf{M}_s^n}{1-F(s)} + \int_0^t \frac{\mathbf{M}_s^n}{[1-F(s)]^2} dF(s) \right\} \\
&= X_t^{n,\alpha} + \frac{1}{n^{\frac{1}{2}-\alpha}} [1-F(t)] \int_0^t \frac{\mathbf{M}_s^n}{[1-F(s)]^2} dF(s).
\end{aligned}
$$

(iv) $\Psi(x_{[0,t]})$ is nothing but **(iii)** with x_t replaced by $\frac{1}{n^{\frac{1}{2}-\alpha}} \mathbf{M}_t^n$. A desired continuity of Ψ follows from

$$
\sup_{t \in [0,1]} |x'_t - x''_t| \le \varepsilon \ \Rightarrow \ \sup_{t \in [0,1]} \left| \Psi\big(x'_{[0,t]}\big) - \Psi\big(x''_{[0,t]}\big) \right| \le 2\varepsilon.
$$

□

2.4.1. **An Alternative Proof of Billingsley's Theorem.** For $\alpha = \frac{1}{2}$, write $X_t^{n,\frac{1}{2}} = \sqrt{n}\big[F_n(t) - F(t)\big]$.

Lemma 2.2. *The family* $\big\{\big(X_t^{n,\frac{1}{2}}\big)_{t\in[0,1]}\big)\big\}_{n\to\infty}$ *converges in law (in Skorokhod's and uniform metrics) to the continuous Gaussian process* $(X_t)_{t\in[0,1]}$ *defined in* (1.2).

Proof. By Lemma 2.1**(iv)**, $X_t^{n,\frac{1}{2}} = \Psi\big(\frac{1}{\sqrt{n}} \mathbf{M}_{[0,t]}^n\big)$ and by Theorem 2.1,

$$
\big(X_t^{n,\frac{1}{2}}\big)_{t\in[0,1]} \xrightarrow[n\to\infty]{\text{law}} \Psi\big(\mathbf{M}_{[0,t]}\big)_{t\in[0,1]}.
$$

Now, by applying the Itô formula to $X_t := \Psi\big(\mathbf{M}_{[0,t]}\big)$, we make sure that X_t solves (1.2). □

3. The Kolmogorov bound

In this section, we show that

$$\begin{aligned}&\mathsf{P}\Big(\sup_{t\in[0,1]}|F_n(t)-F(t)|\ge\varepsilon\Big)\\ &\le 2\exp\bigg(-n\Big[\frac{\varepsilon}{8}\Big\{\log\Big(1+\frac{\varepsilon^2}{32}\Big)-1\Big\}+\frac{4}{\varepsilon}\log\Big(1+\frac{\varepsilon^2}{32}\Big)\Big]\bigg).\end{aligned}\tag{3.1}$$

Since $F_n(t)$ and $F(t)$ are increasing functions and $F(t)$ is continuous, the following upper bound with a free parameter $T\in(0,1)$ holds:

$$\begin{aligned}\sup_{t\in[0,1]}|F_n(t)-F(t)| &\le \sup_{t\in[0,T]}|F_n(t)-F(t)|+\sup_{t\in(T,1]}|F_n(t)-F(t)|\\ &\le \sup_{t\in[0,T]}|F_n(t)-F(t)|+|1-F(T)|+|1-F_n(T)|\\ &\le \sup_{t\in[0,T]}|F_n(t)-F(t)|+2[1-F(T)]+|F(T)-F_n(T)|\\ &\le 2\Big\{\sup_{t\in[0,T]}|F_n(t)-F(t)|+[1-F(T)]\Big\}.\end{aligned}$$

A choice of T with $1-F(T)=\frac{\varepsilon}{4}$ guarantees a useful upper bound

$$\mathsf{P}\Big(\sup_{t\in[0,1]}|F_n(t)-F(t)|\ge\varepsilon\Big)\le\mathsf{P}\Big(\sup_{t\in[0,T]}|F_n(t)-F(t)|\ge\frac{\varepsilon}{4}\Big).$$

By Lemma 2.1 (**(iv)**) with $\alpha=0$, we find that

$$\begin{aligned}\sup_{t\in[0,T]}|F_n(t)-F(t)|&=\sup_{t\in[0,T]}|X^{n,0}_t|\\ &\le\frac{1}{\sqrt{n}}\sup_{t\in[0,T]}|\mathbf{M}^n_t|\Big(1+\sup_{t\in[0,1]}[1-F(t)]\int_0^t\frac{dF(s)}{[1-F(s)]^2}\Big)\\ &\le\frac{2}{\sqrt{n}}\sup_{t\in[0,T]}|\mathbf{M}^n_t|\end{aligned}$$

and the following upper bound:

$$\mathsf{P}\big(\sup_{t\in[0,1]}|F_n(t)-F(t)|\ge\varepsilon\big)\le\mathsf{P}\big(\sup_{t\in[0,T]}\frac{1}{\sqrt{n}}|\mathbf{M}^n_t|\ge\frac{\varepsilon}{8}\big).$$

Now, we shall combine "exponential martingale" and "stopping time" techniques. With $\lambda>0$, let us introduce the exponential martingale

$$\mathfrak{z}_t=\exp\bigg(\frac{\lambda}{\sqrt{n}}\mathbf{M}^n_t-\Big[e^{\frac{\lambda}{n}}-\frac{\lambda}{n}-1\Big]\mathbf{A}^n_t\bigg)\tag{3.2}$$

relative to the filtration $(\mathscr{F}_t)_{t\in[0,1]}$. It is well known that any exponential martingale is a supermartingale too, that is, $(\mathfrak{z}_t,\mathscr{F}_t)_{t\in[0,1]}$ is the nonnegative supermartingale with $\mathsf{E}\mathfrak{z}_\tau\le\mathsf{E}\mathfrak{z}_0=1$ for any stopping time τ w.r.t. the filtration $(\mathscr{F}_t)_{[t\in[0,1]]}$.

We choose two stopping times,

$$\tau^n_\pm=\inf\Big\{t\le T:\pm\frac{1}{\sqrt{n}}\mathbf{M}^n_t\ge\frac{\varepsilon}{8}\Big\},\quad \inf(\varnothing)=\infty,$$

and use them for obtaining the following bound:

$$\mathsf{P}\Big(\sup_{t\in[0,T]}\frac{1}{\sqrt{n}}|\mathbf{M}^n_t|>\frac{\varepsilon}{8}\Big)\le 2\max\Big[\mathsf{P}\Big(\tau^n_+<\infty\Big),\mathsf{P}\Big(\tau^n_-<\infty\Big)\Big].$$

In order to find an upper bound of $\mathsf{P}(\tau_+<\infty)$, write

$$\begin{aligned}1\ge \mathsf{E}\mathfrak{z}_{\tau_+}\ge \mathsf{E}I_{\{\tau_+<\infty\}}\mathfrak{z}_{\tau_+}&=\mathsf{E}I_{\{\tau_+<\infty\}}\exp\Big(\lambda\frac{1}{\sqrt{n}}\mathbf{M}^n_{\tau_+}-\Big[e^{\frac{\lambda}{n}}-1-\frac{\lambda}{n}\Big]\mathbf{A}^n_{\tau_+}\Big)\\ &\ge \mathsf{P}(\tau_+<\infty)\exp\Big(\lambda\frac{\varepsilon}{8}-\Big[e^{\frac{\lambda}{n}}-1-\frac{\lambda}{n}\Big]\mathbf{A}^n_T\Big).\end{aligned}$$

By (2.1), $\mathbf{A}^n_T\le\frac{n}{1-F(T)}=\frac{4n}{\varepsilon}$, so that

$$1\ge \mathsf{P}(\tau^n_+<\infty)\exp\Big(\lambda\frac{\varepsilon}{8}-\Big[e^{\frac{\lambda}{n}}-1-\frac{\lambda}{n}\Big]\frac{4n}{\varepsilon}\Big)$$

or, equivalently, $\mathsf{P}(\tau^n_+<\infty)\le\exp\big(-\big\{\lambda\frac{\varepsilon}{8}-\big[e^{\frac{\lambda}{n}}-1-\frac{\lambda}{n}\big]\frac{4n}{\varepsilon}\big\}\big)$. Since λ is an arbitrary positive parameter, we can set λ as $\lambda^*=\operatorname{argmax}_{\mu>0}\big\{\mu\frac{\varepsilon}{8}-\big[e^{\frac{\mu}{n}}-1-\frac{\mu}{n}\big]\frac{4n}{\varepsilon}\big\}=n\log\big(1+\frac{\varepsilon^2}{32}\big)$, in order to obtain

$$\begin{aligned}&\mathsf{P}(\tau^n_+<\infty)\\ &\le\exp\Big(-\Big\{\lambda^*\frac{\varepsilon}{8}-\Big[e^{\frac{\lambda^*}{n}}-1-\frac{\lambda^*}{n}\Big]\frac{4n}{\varepsilon}\Big\}\Big)\\ &=\exp\Big(-n\Big[\frac{\varepsilon}{8}\Big\{\log\Big(1+\frac{\varepsilon^2}{32}\Big)-1\Big\}+\frac{4}{\varepsilon}\log\Big(1+\frac{\varepsilon^2}{32}\Big)\Big]\Big).\end{aligned}$$

The proof of the upper bound $\mathsf{P}(\tau^n_-<\infty)\le\exp\big(-n\big[\frac{\varepsilon}{8}\big\{\log\big(1+\frac{\varepsilon^2}{32}\big)-1\big\}+\frac{4}{\varepsilon}\log\big(1+\frac{\varepsilon^2}{32}\big)\big]\big)$ is similar.

Therefore, (3) holds. □

4. **The proof of (1.4)**

Recall that $X^{n,\alpha}_t=n^\alpha\big[F_n(t)-F(t)\big]$ (see (2.2)).

Theorem 4.1. *For any $\alpha\in\big(0,\frac12\big)$ and any T in a small vicinity of $\{1\}$,*

$$\lim_{n\to\infty}\frac{1}{n^{1-2\alpha}}\log\mathsf{P}\Big(\sup_{t\in[0,T]}n^\alpha|X^{n,\alpha}_t|\ge\varepsilon\Big)=-2\varepsilon^2.$$

Proof. By Theorem A.2 (Appendix) the family $\{(n^\alpha X^{n,\alpha}_t))_{t\in[0,T]}\}_{n\to\infty}$ obeys the large deviations principle in the Skorokhod space $\mathbb{D}_{[0,1]}$ relative Skorokhod's and uniform metrics with the rate $\frac{1}{n^{1-2\alpha}}$ and the rate function

$$J_T(u)=\frac12\begin{cases}\displaystyle\int_0^T\Big(\dot u_t+\frac{u_t}{1-F(t)}\Big)^2dF(t), & \substack{u_0=0\\ du_t=\dot u_t dF(t)\\ \int_0^T(\dot u_t+\frac{u_t}{1-F(t)})^2dF(t)<\infty}\\ \infty, & \text{otherwise.}\end{cases}$$

Since paths of $(X^{n,\alpha})_{t\in[0,T]}$ with property $\big\{\sup_{t\in[0,T]}n^\alpha|X^{n,\alpha}_t|\ge\varepsilon\big\}$ form a closed set

$$\mathsf{C}=\left\{u\in\mathbb{D}_{[0,T]}:\substack{u_0=0\\ \theta(u)=\inf\{t\le T:|u_t|\ge\varepsilon\}\le T\\ u_t\equiv0,\ t>\theta(u)}\right\},$$

in accordance with the large deviations theory,

$$\varlimsup_{n\to\infty}\frac{1}{n^{1-2\alpha}}\log \mathsf{P}\Big(\sup_{t\in[0,T]} n^\alpha|X^{n,\alpha}_t|\ge\varepsilon\Big)\le -\inf_{u\in\mathsf{C}}J_T(u).$$

A minimization procedure of $J_T(u)$ in $u\in\mathsf{C}$ automatically excludes from consideration all functions $(u_t)_{t\in[0,T]}$ with $J_T(u)=\infty$. Consequently,

$$\varlimsup_{n\to\infty}\frac{1}{n^{1-2\alpha}}\log \mathsf{P}\Big(\sup_{t\in[0,T]} n^\alpha|X^{n,\alpha}_t|\ge\varepsilon\Big)\le -\frac12\inf_{u\in\mathsf{C}}\int_0^{T\wedge\theta(u)}\Big[\dot u_s+\frac{u_s}{1-F(s)}\Big]^2dF(s).$$

Denote $w_t=\dot u_t+\frac{u_t}{1-F(t)}$. Then $J_{\theta(u)}(u)=\frac12\int_0^{\theta(u)}w^2_tdF(t)$, and

$$u_t=-\int_0^t\frac{u_s}{1-F(s)}dF(s)+\int_0^t w_sdF(s),\ t\le\theta(u).$$

This integral equation obeys the unique solution

$$u_{t\wedge\theta(u)}=[1-F(t\wedge\theta(u))]\int_0^{t\wedge\theta(u)}\frac{w_s}{1-F(s)}dF(s).$$

The assumption $\theta(u)\le T$ implies $u^2_{\theta(u)}=\varepsilon^2$. Hence

$$\varepsilon^2=[1-F(\theta(u))]^2\Big(\int_0^{\theta(u)}\frac{w_s}{1-F(s)}dF(s)\Big)^2. \tag{4.1}$$

Now, we apply the Cauchy-Schwarz's inequality,

$$\begin{aligned}\Big(\int_0^{\theta(u)}\frac{w_s}{1-F(s)}dF(s)\Big)^2&\le\int_0^{\theta(u)}\frac{dF(s)}{[1-F(s)]^2}\int_0^{\theta(u)}w^2_sdF(s) \qquad (4.2)\\ &=\frac{F(\theta(u))}{1-F(\theta(u)}2J_{\theta(u)}(u),\end{aligned}$$

transforming (4.1) into the lower bound: $J_{\theta(u)}(u)\ge\frac{\varepsilon^2}{2F(\theta(u))[1-F(\theta(u))]}$. Assume for a moment that there exists u^*_t such that $F(\theta(u^*))=\frac12$. Then the following lower bound $J_{\theta(u^*)}(u^*)\ge 2\varepsilon^2$ is valid. This lower bound is attainable, $J_{\theta(u^*)}(u^*)=2\varepsilon^2$, provided that the Cauchy-Schwarz's inequality in (4.2) becomes the equality. The latter holds true if w^*_s, related to $u^*_t(\dot u^*_t)$, is in a proportion to $\frac{1}{1-F(s)}$, i.e. $w^*_s=\frac{l}{1-F(s)}$ and there exists a constant l^* such that $\big(\int_0^{\theta(u^*)}\frac{w^*_s}{1-F(s)}dF(s)\big)^2=4\varepsilon^2$. The existence of $l^*=2\varepsilon$ is verified directly.

Thus, the upper bound is valid:

$$\varlimsup_{n\to\infty}\frac{1}{n^{1-2\alpha}}\log \mathsf{P}\Big(\sup_{t\in[0,T]} n^\alpha|X^{n,\alpha}_t|\ge\varepsilon\Big)\le -2\varepsilon^2.$$

In order to complete the proof, we have to prove the following lower bound

$$\varliminf_{n\to\infty}\frac{1}{n^{1-2\alpha}}\log \mathsf{P}\Big(\sup_{t\in[0,T]} n^\alpha|X^{n,\alpha}_t|\ge\varepsilon\Big)\ge -2\varepsilon^2$$

Formally, one may apply

$$\varliminf_{n\to\infty}\frac{1}{n^{1-2\alpha}}\log \mathsf{P}\Big(\sup_{t\in[0,T]} n^{\alpha}|X^{n,\alpha}_t|\ge\varepsilon\Big)$$
$$\ge -\frac{1}{2}\inf_{u\in\mathsf{C}^\circ}\int_0^{T\wedge\theta(u)}\Big[\dot u_s+\frac{u_s}{1-F(s)}\Big]^2dF(s),$$

where C° is an interior of C. However, C has an empty interior. Fortunately, the proof of the upper bound gives us a hint: $F(\theta(u^*)=\frac{1}{2}$. Choose T^* with $F(T^*)=\frac{1}{2}$ and use an obvious inequality:

$$\mathsf{P}\Big(\sup_{t\in[0,T]} n^{\alpha}|X^{n,\alpha}_t|\ge\varepsilon\Big)\ge \mathsf{P}\Big(n^{\alpha}|X^{n,\alpha}_{T^*}|\ge\varepsilon\Big).$$

Hence, only a lower bound $\varliminf_{n\to\infty}\frac{1}{n^{1-2\alpha}}\log\mathsf{P}\big(n^{\alpha}|X^{n,\alpha}_{T^*}|\ge\varepsilon\big)\ge -2\varepsilon^2$ has to be proven. The latter is verified with the help of the large deviations principle for the different family $\{X^{n,\alpha}_{T^*}\}_{n\to\infty}$.

Since $X^{n,\alpha}_{T^*}=\frac{1}{n^{1/2-\alpha}}\frac{1}{\sqrt{n}}\sum\limits_{k=1}^n\big[I_{\{\xi_k\le T^*\}}-F(T^*)\big]$ with $(I_{\{\xi_k\le T^*\}}-F(T^*))_{k\ge1}$ being the i.i.d. sequence of zero mean random variables having the variance $F(T^*)[1-F(T^*)]=\frac{1}{4}$, the large deviations principle for this family is well known and has the rate $\frac{1}{n^{1-2\alpha}}$ and the rate function $I(v)=\frac{v^2}{2F(T^*)[1-F(T^*)]}=2v^2$. Therefore,

$$\varliminf_{n\to\infty}\frac{1}{n^{1-2\alpha}}\log\mathsf{P}\Big(n^{\alpha}|X^{n,\alpha}_{T^*}|\ge\varepsilon\Big)=-\inf_{v:|v|\ge\varepsilon}I(v)=-2\varepsilon^2.$$

□

Appendix A. Large deviations principle for $X^{n,\alpha}$

By (2.2), $X^{n,\alpha}_t=-\int_0^t\frac{X^{n,\alpha}_s}{1-F(s)}dF(s)+\frac{1}{n^{\frac12-\alpha}}\mathbf{M}^n_t$. A complicated structure of the martingale $(\mathbf{M}^n_t,\mathscr{F}_t)_{t\in[0,1]}$ does not allow us to apply Freidlin and Wentzell's (1984), [4], or of Wentzell's (1986) [11] results.

On the other hand, by Theorem 2.1, the family $\{(\mathbf{M}^n_t)_{t\in[0,1]}\}_{n\to\infty}$ converges in law to Gaussian martingale $(\mathbf{M}_t)_{t\in[0,1]}$ with $\langle\mathbf{M}\rangle_t=F(t)$. Notice also that the family $\{(\frac{1}{n^{\frac12-\alpha}}\mathbf{M}_t)_{t\in[0,1]}\}_{n\to\infty}$ is in a framework of Freidlin and Wentzell (1984). So, it obeys the large deviations principle with the rate $\frac{1}{n^{1-2\alpha}}$ and the rate function

$$I(u)=\frac{1}{2}\begin{cases}\int_0^T\dot u_t^2dF(t), & {}^{u_0=0}_{\substack{du_t=\dot u_tdF(t)\\ \int_0^T\dot u_t^2dF(t)<\infty}}\\ \infty, & \text{otherwise.}\end{cases}\tag{A.1}$$

Theorem A.1. *For any $\alpha\in\big(0,\frac12\big)$ and any T in a small vicinity of $\{1\}$, the families*

$$\Big\{\Big(\frac{1}{n^{\frac12-\alpha}}\mathbf{M}_t\Big)_{t\in[0,T]}\Big\}_{n\to\infty}\quad\text{and}\quad\Big\{\Big(\frac{1}{n^{\frac12-\alpha}}\mathbf{M}^n_t\Big)_{t\in[0,T]}\Big\}_{n\to\infty}$$

share the same large deviations principle.

Proof. Hereafter, $\lambda(s)$ is a bounded measurable function.

Since $\frac{1}{n^{\frac12-\alpha}}\int_0^t\lambda(s)d\mathbf{M}_s$ is a continuous Gaussian martingale with the predictable variation process $\frac{1}{n^{1-2\alpha}}\int_0^t\lambda^2(s)dF(s)\ =:\ 2\widetilde{\mathscr{E}}^{n,\alpha}_t(\lambda)$, the function $\widetilde{\mathscr{E}}^{n,\alpha}_t(\lambda)$ is the Laplace transform of $\frac{1}{n^{\frac12-\alpha}}\int_0^t\lambda(s)d\mathbf{M}_s$. Moreover, a random process $\widetilde{\mathfrak{z}}^{n,\alpha}_t = \exp\Big(\frac{1}{n^{\frac12-\alpha}}\int_0^t\lambda(s)d\mathbf{M}_s-\log\widetilde{\mathscr{E}}^{n,\alpha}_t(\lambda)\Big)$ is a martingale. In the case of $\frac{1}{n^{\frac12-\alpha}}\int_0^t\lambda(s)d\mathbf{M}^n_s$, an explicit formula for the Laplace transform is unknown. However, a random process $\mathscr{E}^{n,\alpha}_t(\lambda)=\exp\Big(\int_0^t\big[e^{\frac{\lambda(s)}{n^{1-\alpha}}}-1-\frac{\lambda(s)}{n^{1-\alpha}}\big]d\mathbf{A}^n_s\Big)$ "exponentially compensates" $\frac{1}{n^{\frac12-\alpha}}\int_0^t\lambda(s)d\mathbf{M}^n_s$ up to a martingale in a sense that a random process $\mathfrak{z}^{n,\alpha}_t=\exp\Big(\frac{1}{n^{\frac12-\alpha}}\int_0^t\lambda(s)d\mathbf{M}^n_s-\mathscr{E}^{n,\alpha}_t(\lambda)\Big)$ is a local martingale (the latter is verified by applying the Itô formula).

By a terminology of Puhalskii (1994, 2001), $\widetilde{\mathscr{E}}^{n,\alpha}_t(\lambda)$ and $\mathscr{E}^{n,\alpha}_t(\lambda)$ are referred to as "Stochastic Exponentials" related to the families

$$\Big\{\Big(\frac{1}{n^{\frac12-\alpha}}\mathbf{M}_t\Big)_{t\in[0,T]}\Big\}_{n\to\infty}\quad\text{and}\quad\Big\{\Big(\frac{1}{n^{\frac12-\alpha}}\mathbf{M}^n_t\Big)_{t\in[0,T]}\Big\}_{n\to\infty}$$

respectively.

A role of stochastic exponential is revealed in Puhalskii (1994, 2001). In our setting the Puhalskii result states that the above-mentioned families share the same large deviations principle provided that for any $\eta>0$ and any bounded $\lambda(t)$,

$$\lim_{n\to 0}\frac{1}{n^{1-2\alpha}}\log\mathsf{P}\Big(\sup_{t\in[0,T]}n^{1-2\alpha}\Big|\log\frac{\mathscr{E}^{n,\alpha}_t(\lambda)}{\widetilde{\mathscr{E}}^{n,\alpha}_t(\lambda)}\Big|>\eta\Big)=-\infty. \tag{A.2}$$

We finish the proof by verification of (A.2). Taking into account

$$d\mathbf{A}^n_s=n\frac{1-F_n(s)}{1-F(s)}dF(s)$$

(see (2.1)), write

$$\begin{aligned}
&n^{1-2\alpha}\Big|\log\frac{\mathscr{E}^{n,\alpha}_t(\lambda)}{\widetilde{\mathscr{E}}^{n,\alpha}_t(\lambda)}\Big|\\
&=n^{1-2\alpha}\Big|\int_0^t\Big[e^{\frac{\lambda(s)}{n^{1-\alpha}}}-1-\frac{\lambda(s)}{n^{1-\alpha}}\Big]d\mathbf{A}^n_s-\int_0^t\frac{\lambda^2(s)}{2n^{1-2\alpha}}dF(s)\Big|\\
&=n^{1-2\alpha}\Big|\int_0^t\Big[e^{\frac{\lambda(s)}{n^{1-\alpha}}}-1-\frac{\lambda(s)}{n^{1-\alpha}}-\frac{\lambda^2(s)}{2n^{2(1-\alpha)}}\Big]n\frac{1-F_n(s)}{1-F(s)}dF(s)\\
&\quad-\int_0^t\Big[\frac{\lambda^2(s)}{2n^{2(1-\alpha)}}n\frac{1-F_n(s)}{1-F(s)}-\frac{\lambda^2(s)}{2n^{1-2\alpha}}\Big]dF(s)\Big|.
\end{aligned}$$

Since

$$\int_0^T n^{1-2\alpha}\Big|e^{\frac{\lambda(s)}{n^{1-\alpha}}}-1-\frac{\lambda(s)}{n^{1-\alpha}}-\frac{\lambda^2(s)}{2n^{2(1-\alpha)}}\Big|n\frac{1-F_n(s)}{1-F(s)}dF(s)$$
$$\le \text{const.}\frac{1}{n^{1-\alpha}}\int_0^T\frac{1}{1-F(s)}dF(s)=\text{const.}\frac{1}{n^{1-\alpha}}\log\frac{1}{1-F(T)}$$

and

$$\int_0^T n^{1-2\alpha}\Big|\frac{\lambda^2(s)}{2n^{2(1-\alpha)}}n\frac{1-F_n(s)}{1-F(s)}-\frac{\lambda^2(s)}{2n^{1-2\alpha}}\Big|dF(s)$$
$$\le\int_0^T\frac{\lambda^2(s)}{2}\frac{|F_n(s)-F(s)|}{1-F(s)}dF(s)$$
$$\le\text{const.}\sup_{s\in[0,T]}|F_n(s)-F(s)|\int_0^T\frac{1}{1-F(s)}dF(s)$$
$$=\text{const.}\sup_{s\in[0,T]}|F_n(s)-F(s)|\log\frac{1}{1-F(T)},$$

we shall analyze an upper bound of the following inequality:

$$n^{1-2\alpha}\Big|\sup_{t\in[0,T]}\log\frac{\mathscr{E}_t^{n,\alpha}(\lambda)}{\widetilde{\mathscr{E}}_t^{n,\alpha}(\lambda)}\Big|$$
$$\le\text{const.}\log\frac{1}{1-F(T)}\Big[\frac{1}{n^{1-\alpha}}+\sup_{s\in[0,T]}|F_n(s)-F(s)|\Big].$$

Obviously, (A.2) is valid if

$$\lim_{n\to\infty}\frac{1}{n^{1-2\alpha}}\mathsf{P}\big(\sup_{s\in[0,T]}|F_n(s)-F(s)|>\eta-\frac{1}{n^{1-\alpha}}\big)=-\infty.$$

For fixed η, let us choose a number n_0 such that $\frac{1}{n_0^{1-\alpha}}\le\frac{\eta}{2}$ and all $n\ge n_0$. In this scenario it remains to show that

$$\lim_{n\to\infty}\frac{1}{n^{1-2\alpha}}\mathsf{P}\Big(\sup_{s\in[0,T]}|F_n(s)-F(s)|>\frac{\eta}{2}\Big)=-\infty.$$

The latter heavily uses Kolmogorov's bound:

$$\frac{1}{n^{1-2\alpha}}\log\mathsf{P}\bigg(\sup_{s\in[0,T]}|F_n(s)-F(s)|\ge\frac{\eta}{2}\bigg)$$
$$\le\frac{\log 2}{n^{1-2\alpha}}-n^{2\alpha}\Big[\frac{\eta}{16}\Big\{\log\Big(1+\frac{\eta^2}{128}\Big)-1\Big\}+\frac{8}{\eta}\log\Big(1+\frac{\eta^2}{128}\Big)\Big]\bigg)\xrightarrow[n\to\infty]{}-\infty.$$

$\square$

Theorem A.1 implies the following result.

Theorem A.2. *For any $\alpha\in\big(0,\frac{1}{2}\big)$ and any T in a small vicinity of $\{1\}$, the family $\{(X_t^{n,\alpha})_{t\in[0,T]}\}_{n\to\infty}$ obeys the large deviations principle in the Skorokhod space $\mathbb{D}_{[0,T]}$ relative Skorokhod's and uniform metrics with the*

rate speed $\frac{1}{n^{1-2\alpha}}$ *and the rate function*

$$J_T(u) = \frac{1}{2}\begin{cases} \int_0^T \Big(\dot u_t + \frac{u_t}{1-F(t)}\Big)^2 dt, & \substack{u_0=0 \\ du_t=\dot u_t dF(t) \\ \int_0^T (\dot u_t + \frac{u_t}{1-F(t)})^2 dF(t)<\infty} \\ \infty, & \textit{otherwise}. \end{cases}$$

Proof. By Lemma 2.1, $X^{n,\alpha}_t = \Psi\Big(\frac{1}{n^{\frac12-\alpha}}\mathbf{M}^n_{[0,t]}\Big)$. Hence and by Theorem A.1 the family $\{(X^{n,\alpha}_t)_{t\in[0,T]}\}_{n\to\infty}$ shares the large deviations principle with the family $\Psi\Big(\frac{1}{n^{\frac12-\alpha}}\mathbf{M}_{[0,t]}\Big)_{t\in[0,T]}$.

Hence, by the contraction principle of Varadhan (1984) and (A.1) $J_T(u) = I(v)_{v=\Psi(u)}$. □

References

[1] Billingsley, P. (1968). *Converges of probability measures*. Wiley.

[2] Dembo, A., Zeitouni, O. (1998). *Large Deviations Techniques and Applications*. Springer, 2nd edition.

[3] Doob, J. L. (1953). *Stochastic Processes*. N.Y., John Wiley and Sons, London-Chapman and Hall

[4] Freidlin, M.I., Wentzell A.D. (1984). *Random Perturbations of Dynamical Systems.* N.Y. Springer.

[5] Liptser, R. Sh. and Shiryayev, A. N. (1989). Theory of martingales. Mathematics and its Applications (Soviet Series), 49. Kluwer Academic Publishers Group, Dordrecht.

[6] Liptser, R. Sh. and Shiryaev, A. N. (2000). *Statistics of Random Processes II*, 2nd ed., Springer, Berlin - New York.

[7] Puhalskii, A.A. (1994). The method of stochastic exponentials for large deviations. *Stochast. Proc. Appl.* **54**, , pp. 45-70.

[8] Puhalskii, A. (2001). Large Deviations and Idempotent Probability, Chapman & Hall/CRC Press.

[9] Sanov, I.N. (1961). On the probability of large deviations of random variables (in Russian). (English translation from *Math. Sb.*(42)). In selected Translation in Mathematics, Statiscs and Probability, I (1961), 213-244.

[10] Varadhan, S.R.S. (1984) *Large Deviations and Applications.* SIAM, Philadelphia.

[11] Wentzell, A.D. (1986). *Limit Theorem on Large Deviations for Markov Stochastic Processes.* Nauka (in Russian, English translation: Kluwer, 1990).

[12] Wu Liming (1994). Large Deviations, Moderate Deviations and LIL for Empirical Provesses. *The Annals of Probability.*

Department of Electrical Engineering Systems, Tel Aviv University, 69978 Tel Aviv, Israel

E-mail address: liptser@eng.tau.ac.il; rliptser@gmail.com